Check for updates

Clinical interpretation of an interpretable prognostic model for patients with COVID-19

Daniele Roberto Giacobbe ✉

Li Yan and colleagues described an interesting predictive model of unfavourable prognosis in critically ill patients with coronavirus disease 2019 (COVID-19) in a single-centre study in China[1]. The model was developed and validated using interpretable machine learning techniques (in this case, a supervised XGBoost classifier modified into a 'single-tree XGBoost' to further increase interpretability)[1]. The model selected three biomarkers (high-sensitivity C-reactive protein (hs-CRP), lactic dehydrogenase (LDH) and lymphocyte counts) that could be used to predict mortality more than 10 days in advance with >90% accuracy.

Along with interpretability of the model, interpretation of clinical data remains paramount. In this regard, the authors correctly reported that persistent increases in hs-CRP levels reflect persistent inflammation and may be an important marker for poor prognosis in acute respiratory distress syndrome[1]. However, very high hs-CRP levels may also reflect the development of severe bacterial superinfection, which is expected to be a frequent complication in critically ill patients with COVID-19, and potentially the true reason for increased mortality in a non-negligible proportion of them[2,3]. In addition, LDH levels may increase and possibly predict prognosis in severe bacterial infections[4,5]. Although all of this may not strictly change the interpretation of the performance of the model (patients who develop bacterial superinfections are in fact at higher risk of mortality), it may change the clinical perspective, shifting the interest of clinicians from identifying patients at higher risk of death to identifying those at higher risk of superinfections or other complications (to prevent or detect them earlier) before observing substantial increases in hs-CRP and LDH levels. This is just a possible alternative, non-mutually exclusive explanation of the results

of the model by Li Yan and colleagues, which nonetheless stress the importance of a continuous, proactive dialogue between clinicians and developers of machine learning models to fuel further progress in this area.

References

1. Yan, L. et al. An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.* **2**, 283–288 (2020).
2. Cox, M. J., Loman, N., Bogaert, D. & O'Grady, J. Co-infections: potentially lethal and unexplored in COVID-19. *Lancet Microbe* **1**, e11 (2020).
3. Markanday, A. Acute phase reactants in infections: evidence-based review and a guide for clinicians. *Open Forum Infect. Dis.* **2**, ofv098 (2015).
4. Lu, J. et al. Lactate dehydrogenase is associated with 28-day mortality in patients with sepsis: a retrospective observational study. *J. Surg. Res.* **228**,314–321 (2018).
5. Zein, J. G., Lee, G. L., Tawk, M., Dabaja, M. & Kinasewitz, G. T. Prognostic significance of elevated serum lactate dehydrogenase (LDH) in patients with severe sepsis. *CHEST* **126**, 873S (2004).

Clinica Malattie Infettive, Ospedale Policlinico San Martino–IRCCS, Genoa, Italy. ✉e-mail: danieleroberto.giacobbe@hsanmartino.it